

# White Paper Report

Report ID: 102634

Application Number: HD5124411

Project Director: Bryan Wagner (bwagner@berkeley.edu)

Institution: University of California, Berkeley

Reporting Period: 8/1/2011-8/31/2012

Report Due: 11/30/2012

Date Submitted: 11/7/2012

# **Final Report**

NEH Startup Grant HD-51244-11

“A Text Analysis Tool for Examining Stylistic Similarities in  
Narrative Collections”

Project Director: Bryan Wagner  
University of California, Berkeley

August 24, 2012

## Part I

# Narrative Description

## 1 Project Activities

This report describes our one-year progress on the WordSeer text analysis system supported by start-up grant HD-51244-11. The project is to be continued under implementation grant number HK-50011.

In our startup grant proposal, we described our goals as follows :

We are applying for a Level II funding to build and evaluate a system that can compare two or more narratives grammatical structures and authorship features. Our goal is for our English scholars to be able to use our system to gather accurate information about stylistic similarity in manner satisfying to them. From this we hope to learn whether computational linguistics and information visualization can have a place among the research methods of today's literature scholars.

We took a case-study based approach to the question. By observing literature scholars studying problems of real interest to them, we created visualizations, computational linguistic tools, and interaction methods that supported their specific analysis needs throughout their study.

The first question we chose concerned a collection of pre-civil war slave narratives – biographies and autobiographies of escaped former slaves. Professor Bryan Wagner (our director and literary scholar) was interested in ‘style, syntax, and representative tropes’ in the slave narratives. For him, this was part of a larger project examining how these factors influenced the social historical arguments made from the narratives. One of his questions was fairly specific: how closely did the slave narratives collection adhere to a list of conventions described in detail by the scholar James Olney? In a 1984 paper entitled “I was born: Slave Narratives and their status as autobiography”, Olney set out a numbered list of narrative stereotypes that, he asserted, were so “early and firmly established” in the slave narratives that they constituted a sort of “master plan” [14]. This plan includes, among other things, descriptions of cruel masters, details of first observed whippings, attempts to escape guided by the North Star, families being separated from each other, etc.

The WordSeer-based analysis of these stereotypes was a success: we presented our findings at an MLA 2012 panel on “Digital Approaches to the Archive of American Slavery” [10]. WordSeer helped uncover evidence (described fully in Appendix C) that Olney’s numbered list of conventions might be incomplete and somewhat inaccurate, and may not occur in a fixed order.

In addition to publications at academic venues, the project received informal attention from the online digital humanities community. The project’s computer scientist wrote articles about the system and the computational techniques she was using on her blog, and posted about computational linguistics and visualization on her twitter feed. The content was written for a digital humanities audience, as a way to engage the community in a discussion around the idea of computational analysis of literature.

Our online, informal engagement the community proved exceptionally valuable. Through conversations at THATCamps and on twitter, our project attracted two more literary collaborators around the six-month mark. We were joined by Shakespeare scholar Michael Ulliot at the University of Calgary, and Stephen Crane scholar Natalia Cecire. Together with Bryan Wagner, our project director, these scholars were sources of expertise and experience in the field of literature study. Their interest gave us the opportunity to study applications of our tools outside the slave

narratives, and to demonstrate the truth of our claims that our software could be applied to a variety of input text collections. We created two more versions of our system, one on the complete works of Shakespeare, and one on electronic texts of Stephen Crane’s work collected by hand from various online sources such as project Gutenberg. The implementation phase of this project will include in-depth case studies of literary analyses of these two collections.

As described in our start-up proposal, our development process consisted of several incremental phases of feature design, implementation, testing, and redesign. In each phase, we implemented features to address our scholars’ current needs. The scholars would attempt to use the tool, revealing problems with the design and also uncovering new analysis desires. Then a new phase would begin: the tool would be redesigned and the process repeated. WordSeer has completed three such phases of development and is now in its fourth phase, which will continue under the implementation grant. A full description of the current system and all its features is given in Appendix A.

The system’s features do not overlap entirely with the system described in the proposal. Grammatical search, visualization, reading, and exploration exist in both, and both are focused on exploring stylistic similarity. However, the current system focuses on similarity in terms of stereotypes, themes, and conventions, whereas the proposed system would have explored similarity in terms of authorial style. This is because at the time, our scholars were interested in similarity in general terms – authorship being one of them, but during the course of their work, they began to find stereotypes and conventions more interesting. We were also joined by scholars of Shakespeare and Stephen Crane who found tools for the study of thematic similarity interesting. The development of the tool therefore proceeded in that direction.

The current system also has many more features, and is more powerful and broad than the system we originally proposed. The proposal was very focused on algorithms and visualizations. However, we found that when navigation, comparison, and exploration features were added, those same algorithms and visualizations yielded more insight and became more useful. We broadened our efforts to include understanding the literary analysis process as a whole, so that we could better understand how to integrate computational methods into it. The project now has a significant research component in the direction of understanding the process of literary study.

Our experiences with combining literature study, data visualization, and computational linguistics allowed us to reach a better understanding of the computational analysis needs of literature scholars. We re-defined the goals of our project. It is no longer *just* a tool for examining stylistic similarities in slave narratives. During the coming implementation-grant phase, it will be an environment for the computational analysis of literary text.

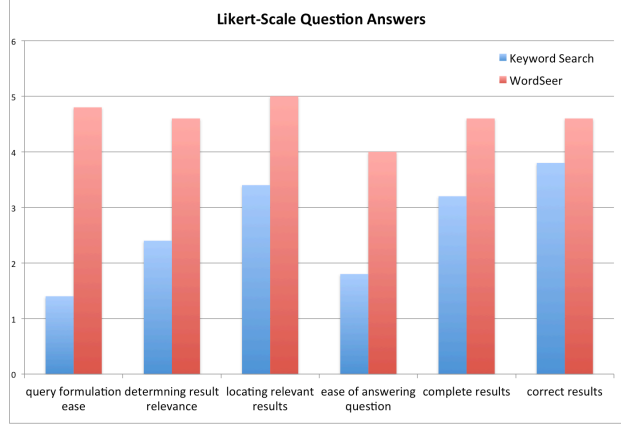
## 2 Timeline of Accomplishments

### 2.1 Phase 1: Search

*September – December 2010*

The highly specific nature of Olney’s stereotypes inspired the first cycle of development. This cycle resulted in the development of WordSeer’s grammatical search feature (described in Appendix A.2), which allows scholars to explore the style and syntax with which concepts are discussed.

As mentioned in our proposal’s “System Description” section, we used algorithms called syntactic parsers (see [3] for an overview) to analyze each sentence and automatically extract relationships between words. Then, a user interface was built to allow queries over word-relationship pairs. As a consequence, instead of simply searching for keywords, scholars could describe highly specific queries, such as “overseer *described as* cruel”. The reading interface (Appendix A.6) was built to



**Figure 1:** Answers to 7-point-scale questions.

accompany the search interface. When an interesting result was seen, scholars could click on it to read the surrounding text at that point.

### 2.1.1 External Evaluation

*December 2010*

We recruited 5 graduate students from the departments of English and History at UC Berkeley. In a 45-minute study, the participants were first shown a walk-through of the interface, and a standard keyword-based interface. Then, they were given three tasks to be done on each interface while thinking aloud, one easy, one medium, and one hard. Together, the tasks built towards typifying a pre-selected type of event in the narratives. The questions were, “give a list of 5 adjectives that characterize slave punishment”, and “describe 5 places to which slaves escaped”. The easy task was to collect a set of narratives that included the event, the medium task was to identify and select patterns of interest in the texts relevant to the question, and the hard task was to answer a summary question. We gave each user both questions, but alternated the interfaces on which users did the tasks.

We compared the users’ effectiveness, efficiency, and satisfaction on the two interfaces. We asked them 7-point-scale questions about how easy it was for them to formulate their query, how easy it was to determine the relevance of results, and how satisfied they were with the accuracy and completeness of their results. The results are shown in Figure 1. WordSeer made it significantly easier for users to formulate their queries ( $p = 0.01$ ), and to answer the question ( $p = 0.01$ ). It was also easier to determine the relevance of search results, with  $p = 0.07$  (marginally significant). The other dimensions were also positive for WordSeer, but we could not establish significance due to our small sample size.

## 2.2 Phase 2: Visualization and Annotation

*January – September 2011*

By examining examples retrieved through grammatical search, the scholars were able to come up with sets of words and relationships that represented many of Olney’s stereotypes. In the second phase, we devised features to help scholars collect examples and understand their prevalence.

WordSeer’s annotation feature (described in Appendix A.6) was designed to help our scholars collect examples. Using this tool, they could read the texts of the narratives, highlight sections

of text, and tag them or annotate them with the stereotype that they seemed to represent. The newspaper-strip visualization was designed to help examine patterns of occurrence of words. Our scholars were interested in whether stereotypes occurred at the beginnings, ends, or middles of narratives. We therefore designed a visualization of the entire collection in which every narrative was the same length, so that they could spot such patterns (Appendix A.4).

### 2.2.1 Results

*May – September 2011*

At the end of phase 2, we were able to conduct our first complete analysis. The results (Appendix C) were presented at the MLA 2012 convention at a panel entitled “Networks, Maps, and Words: Digital Humanities approaches to the Archive of American Slavery” [10]. Our results suggested that some of Olney’s stereotypes were true, but that his “master plan” was overly specific and yet incomplete. There were indeed certain events – cruel treatment, the destruction of families, escape – that occurred so frequently in the collection as to be rightly called stereotypical (C.1). However there seemed to be other common events, such as the death of parents, that he did not mention (C.2), and other supposedly stereotypical events, such as escapes guided by the North Star, that there were not as frequent as seemed implied (C.3).

### 2.2.2 Disseminated Research

Phase 2 resulted in the publication of two conference papers and several research seminars and talks. The system was presented at the 2011 Digital Humanities conference [6], and the 2011 conference on Human-Computer Interaction and Information Retrieval [11]. We also gave talks at the Berkeley Institute for Design (BIDLab) [7], the Maryland Institute for Technology in the Humanities (MITH) [5], and at two THATCamps.

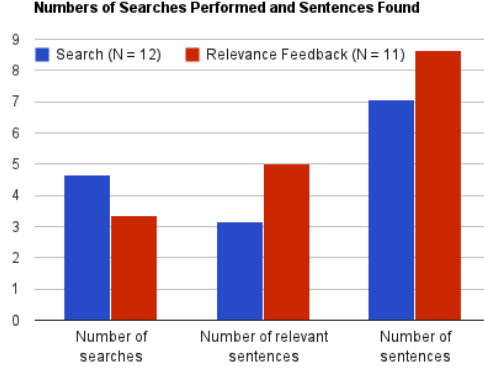
## 2.3 Phase 3: Sensemaking

*October 2011 – present*

Over the course of our phase 2 analysis, it became apparent that WordSeer still had many limitations. Many stereotypes were hard to characterize in terms of grammatical search queries, and we had no way to curate and compare narratives within the collection we were analyzing. For example, there was no way to exclude certain types of documents (such as letters and poems), or certain date ranges, from their analyses, or to compare different subsets. WordSeer is completing phase 3 of development to overcome these problems.

A solution to the data curation and comparison problem was completed in May 2012 – we added a flexible “collections” feature to WordSeer, as described in Appendix A.3. Scholars can collect documents, sentences, passages, and words into collections, which can be moved, reorganized and compared against each other. In previous versions, in order to compare visualizations and grammatical search results, users would have to switch between collections using a drop-down menu. We have developed a free-form desktop-like interface (Appendix B, Figures 15 – 16) which would allow side-by-side comparisons.

We also implemented related words (described in Appendix A.7) as rough solution to the exploration problem. We used computational linguistics algorithms [4] to calculate the similarity between words. When users click on a word while reading, a popup appears showing other words that tend to be used in similar contexts. Related words are also accessible outside the reading interface. Wherever words occur, in search results or visualizations, they can be clicked to display related words.



**Figure 2:** A comparison of the average number of searches performed, number of sentences found, and the number of relevant sentences found across the two experiment conditions in our theme-finding study. Example-based exploration (a.k.a. relevance feedback) is red, and keyword-search only is in blue. The differences are consistent with our hypothesis that example-based exploration (red) makes it easier to find sentences relevant to a theme. With example-based exploration, there are fewer searches, but more sentences, and more relevant sentences.

With the Shakespeare collection available as a result of our collaboration with Dr. Ullyot, we performed a proof-of-concept analysis of the differences in Shakespeare’s descriptions of men and women. This analysis was intended to exploit WordSeer’s new sensemaking capabilities for comparison, curation, and exploration. Using the new features, we were able to uncover evidence of differences in language use across the genders, and across genre. We found that, in plays in which love is a major plot point, the language around women is more physical, and the language around men is more sentimental than in plays that are not primarily about love.

Finally, as a way to help scholars find examples of concepts that are difficult to express in words, we developed a feature for example-based exploration. With example-based exploration, the scholar can specify a few starter examples of a concept that have found using search or visualization. The system examines the examples and uses algorithms for text similarity [16, 15] to retrieve other “similar” sentences. These are presented to the scholar for feedback, and the scholar can mark the sentences as “relevant” or “not relevant”. The algorithm learns from this feedback and repeats the cycle using the updated information. Over multiple feedback cycles from the scholar, it gradually triangulates upon the examples of the concept of interest.

### 2.3.1 Evaluation

Our goal was to determine whether example-based exploration was better able to support finding literary themes than search alone. We decided upon an online study design with a single theme from Shakespeare. Participants were shown an explanation of the theme, with two examples. Then, they were asked to find as many more examples of the same theme as they could within five minutes. Since we were dealing with literary themes, we decided to restrict participation to people with at least college-level backgrounds in English language or literature.

The theme in the study was chosen by our domain-expert collaborator, Dr. Michael Ullyot at the University of Calgary. Once all participants had finished, we submitted their sentences to him. He marked each sentence either relevant or not relevant to the chosen theme. 23 participants with the requisite background completed the study.

Our results were encouraging. The average values in the relevance-feedback condition differed from the average values in the search condition in the direction consistent with our hypothesis. On

average, participants in the relevance feedback condition found more sentences, of which more were relevant, and performed fewer searches in order to do so (Figure 2). However, a bigger study, which we plan to execute in the fall of 2012, is required to establish statistical significance.

### 2.3.2 Disseminated Research

Our proof of concept Shakespeare analysis resulted in a poster presentation at the computer science conference CHI 2012, as an example of a sensemaking interface [13]. We also gave talks at the American Literature Association’s 2012 meeting [9], and the Stanford Human-Computer Interaction group’s weekly meeting [8]. Finally, an article entitled “Supporting Sensemaking in Text Analysis for the Humanities” detailing our Phase 3 efforts was accepted for publication in the journal *Literary and Linguistic Computing*. Lastly, we detailed our investigations into example-based exploration in a paper submitted to the computer science conference HCIR 2012 [12].

In this phase, a news article on WordSeer was also featured on the University of California’s research website [2].

## Bibliography

- [1] S. G Eick. Graphically displaying text. *Computational and Graphical Statistics*, 3(2):127–142, 1994.
- [2] Nicole Freeling. Trimming time in the stacks. <http://universityofcalifornia.edu/research/stories/2011/12/wordseer.html>, December 2011. Accessed January 13th, 2012 2:02 PM.
- [3] Dan Jurafsky and James H. Martin. Chapter 13 syntactic parsing. In *Speech and language processing*, pages 427 — 459. Pearson Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2009.
- [4] Dekang Lin. Using syntactic dependency as local context to resolve word sense ambiguity. In *Proc. 35th annual meeting of the Association for Computational Linguistics*, volume 2, pages 64–71, Stroudsburg, PA, USA, 1997. Association for Computational Linguistics.
- [5] Aditi Muralidharan. Large scale text analysis in the digital humanities: Methods and challenges. MITH Digital Dialogues, <http://mith.umd.edu/podcast/?podcast=112>, September 2011. Accessed January 13th, 2012 1:59 PM.
- [6] Aditi Muralidharan. A visual interface for exploring language use in slave narratives. In *Proc. Digital Humanities 2011*, Stanford, CA, 2011. Stanford University Libraries.
- [7] Aditi Muralidharan. WordSeer: an interface for exploring language use in literary text. Berkeley Institute of Design Seminar <http://www.ischool.berkeley.edu/node/20518>, October 2011. Accessed January 13th, 2012 2:00 PM.
- [8] Aditi Muralidharan. Sensemaking for literature study. Stanford HCI Group Weekly Meeting <http://www.cs.berkeley.edu/~aditi/presentations/sensemaking-for-literature-study.pdf>, 2012.
- [9] Aditi Muralidharan. Text analysis in the (literary) real world: A report on three case studies. Annual Conference of the American Literature Association <http://www.cs.berkeley.edu/~aditi/presentations/ala-2012.pdf>, 2012.



- [10] Aditi Muralidharan. Using digital tools to explore narrative conventions in the north american antebellum slave narratives. Modern Language Association, Session 468. Networks, Maps, and Words: Digital-Humanities Approaches to the Archive of American Slavery, January 2012.
- [11] Aditi Muralidharan and Marti A. Hearst. WordSeer: exploring language use in literary text. In *Proc. HCIR Conference on Human Computer Interaction and Information Retrieval*, Mountain View, CA, 2011.
- [12] Aditi Muralidharan and Marti A. Hearst. Finding literary themes with relevance feedback. In *(submitted) Proc. HCIR 2012* <http://www.cs.berkeley.edu/~aditi/papers/hcir-2012.pdf>, Cambridge, MA, 2012.
- [13] Aditi Muralidharan and Marti A. Hearst. Wordseer: A sensemaking environment for literature study. In *Proc. CHI 2012* <http://www.cs.berkeley.edu/~aditi/papers/chi12-wip-wordseer.pdf>, Austin, TX, 2012.
- [14] James Olney. “i was born”: Slave narratives, their status as autobiography and as literature. *Callaloo*, 20:46–73, January 1984. ArticleType: research-article / Full publication date: Winter, 1984 / Copyright 1984 The Johns Hopkins University Press.
- [15] J. J. Rocchio. Relevance feedback in information retrieval. In Gerard Salton, editor, *The SMART Retrieval System: Experiments in Automatic Document Processing*, Prentice-Hall Series in Automatic Computation, chapter 14, pages 313–323. Prentice-Hall, Englewood Cliffs, NJ, USA, 1971.
- [16] Gerard Salton and Chris Buckley. Improving retrieval performance by relevance feedback. In K.S. Jones and Peter Willett, editors, *Readings in information retrieval*, Morgan Kaufmann series in multimedia information and systems. Morgan Kaufman, 1997.
- [17] eds. TEI Consortium. Guidelines for electronic text encoding and interchange. <http://www.tei-c.org/P5>, 2011.
- [18] Michael Wattenberg and Fernanda B. Viegas. The word tree, an interactive visual concordance. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1221–1228, 2008.

# Part II

## Appendices

### A Current System

In this section, we demonstrate WordSeer’s current capabilities. As a running example, we use the following open-ended question:

“How does the portrayal of men and women in Shakespeare’s plays change under different circumstances?”

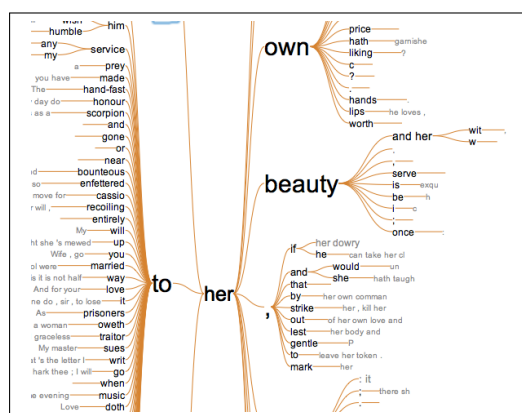
We illustrate how a scholar using WordSeer can find evidence that when love is a major plot point, the language referring to women changes to become more physical, and the language referring to men becomes more sentimental.

#### A.1 Input

The tool is run on a collection of text documents. The input is a set of XML files in a directory, each representing a document in the collection, and the output is a web application with search, visualization, and annotation capabilities. We chose XML because TEI [17], an XML specification for encoding documents, is a widely-adopted digitization standard in the humanities. Many documents of interest to literature scholars are encoded as TEI-XML files.

#### A.2 Search

One might begin the analysis with a question, “what are some things that are portrayed as ‘his’ and some things that are ‘hers’?”. With typical keyword-search systems, the query would be **his** for men, or **her**, and **hers** for women. Such a system would return an unstructured list of results. The word **his** is always a possessive pronoun, so results sequences containing **his** would nearly always be relevant. But **her** can also be a 3rd-person pronoun, and will yield constructions like “I told her that X” and “I gave her the Y”.



**Figure 3:** Word tree [18] for the word “her” generated by WordSeer on Shakespeare’s complete works

and allows users to specify *both* keywords *and* relationships between them. We know of no other text analysis systems in the humanities that apply this technology.

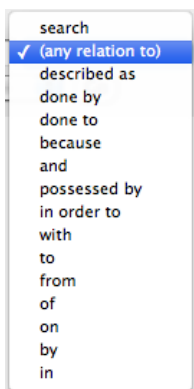
A standard approach in literature study is to view search results in a *concordance*: a list of all the sentences in which a word occurs, with the target word aligned in the center of the view, exposing the contexts to its left and right, sorted in some manner. WordSeer uses the word tree visualization [18] which makes common contexts in a concordance easier to view by grouping them in an arced tree-like structure. The word tree for **her** is shown in Figure 3. Some words like **beauty** stand out, but constructions like **her own** muddy the picture.

With the WordSeer project, we make headway on this problem by providing an easy interface to view the results of *grammatical search* (Figure 4). The system uses natural language processing (NLP) to extract relationships between words (see [3] for an overview),

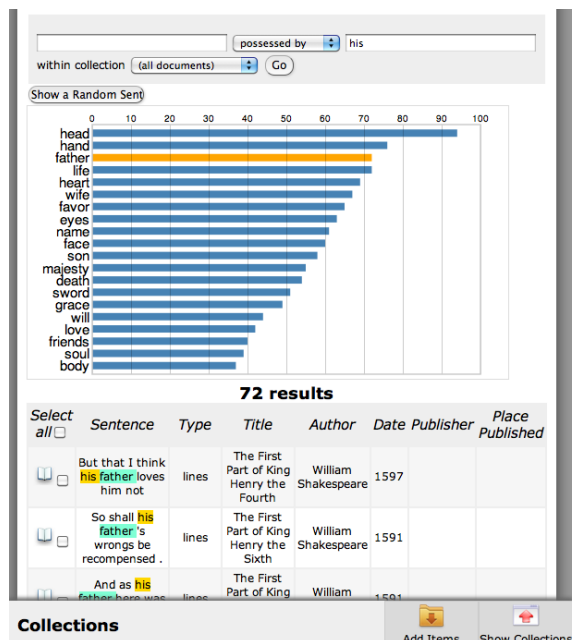
In the tool’s search interface, pairs of words are specified using input boxes, and the relationship between them is selected from a drop-down menu (Figure 5). Leaving a word-input box blank returns all matches.

With this feature, a scholar can, for example, take advantage of the fact that possessive relationships between words can be automatically detected, to express the question precisely: “what are all the words with which **her** has a possessive relationship”.

By comparing the results of various searches, a scholar can begin assemble a picture of language use around a concept. For example, Figure 4 shows search results for all words for which **his** has a possessive relationship. Comparing these words with those for **her** (Figure 6) reveals immediate differences. The word **father** is most common for **her**, with **husband**, and **son** close behind. Several body parts enter the picture: **eyes**, **hand**, **face**, **tongue**, **lips**, **cheek**. A picture emerges: women’s most commonly-mentioned possessions are their male relatives and their bodies.



**Figure 5:** Searchable grammatical relationships between words.



**Figure 4:** Grammatical search results for *possessed-by his*, filtered on the word *father*. Individual search results, corresponding to matching sentences, are highlighted to show the words in the relationship

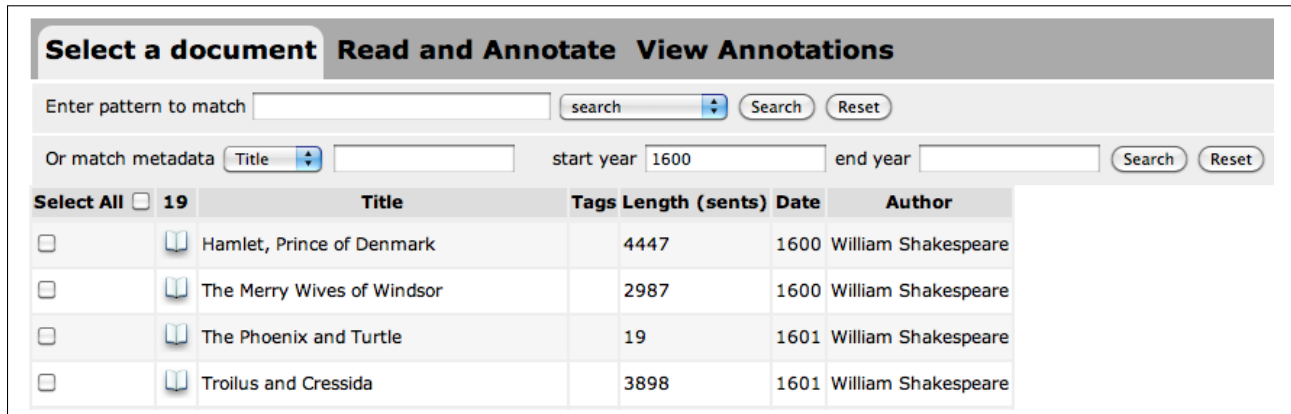
A caveat to such an analysis is that even state-of-the-art language processing algorithms are not 100% accurate. The bigger the disparity between their training data and the input text, the more their accuracy degrades from the reported figures of around 85%.

### A.3 Collection and Comparison

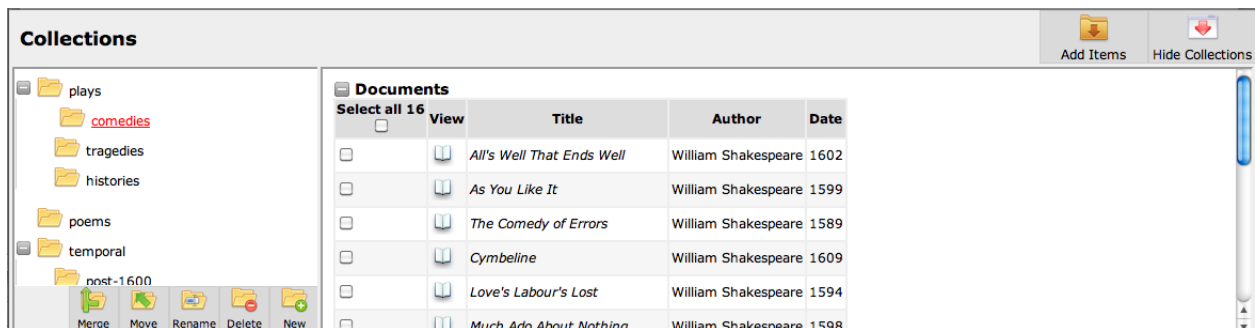
Comparative analysis is central to literature study. WordSeer’s collections feature (Figure 8) supports this mode of analysis by allowing users to collect and organize documents into hierarchical sub-collections. Collections can be moved, renamed, merged, and modified. The document listing (Figure 7) is used to add the appropriate plays, which are sortable and filterable by date, title, full-text search, grammatical search, and length.

For example, upon seeing the physical, patriarchal portrayal of women revealed by grammatical search, a scholar might wish to investigate whether this picture is consistent, or whether it changes in different types of plays. WordSeer’s collections feature can be used to divide the plays into *comedies*, *tragedies*, and *histories* – the three most commonly-accepted categorizations of Shakespeare’s plays. Collections are created using the “collections” bay, a collapsible window at the bottom of the screen. Temporal differences can be investigated by creating the *pre-1600*, and *post-1600* categories.

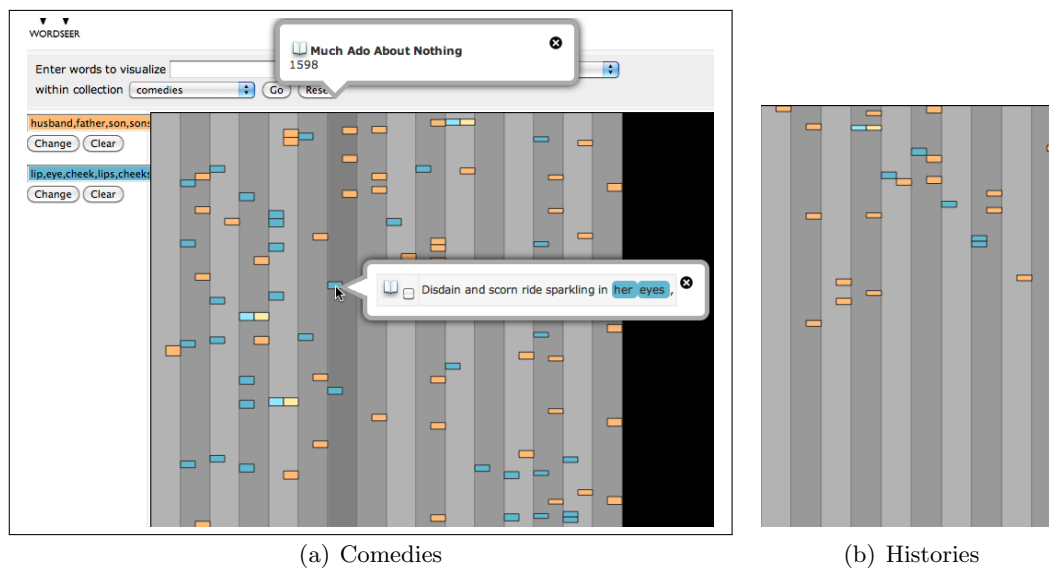
### A.4 Visualization



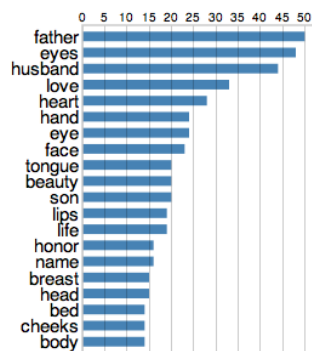
**Figure 7:** The document listing, used to select a document to read, or to add to a collection. Documents can be filtered using a grammatical or keyword search query on the text, by search on title or author, or by date. For easy scanning, the list of results is sortable by title, date, document length, and author.



**Figure 8:** The collections bay showing a number of collections containing documents. Collections can also contain words, sentences, and "snippets" of text .



**Figure 9:** Comparing the prevalence of body parts *possessed-by* her (eyes, lips, cheeks, and face)(blue) and relatives *possessed-by* her (husband, father, sons, daughters, children) (orange) in the comedies and histories. Each column is a play, represented in alternating shades of grey.



**Figure 6:** Results for the grammatical search *possessed-by her*. The bar graph shows the 20 most-frequent matches along with their frequencies.

that body part references are somewhat less prevalent in the historical plays, but family remains unchanged.

A scholar can investigate why body parts seem less prevalent in the historical plays by hovering over a few highlighted blocks corresponding to body-part results (blue). A rough sample will find that many of the sentences in which body parts occur are romantic in nature. The reading and annotating interface can be used to follow up on this phenomenon clicking on the highlighted blocks in the newspaper-column visualization.

## A.5 Close Reading

Hamlet, Prince of Denmark	Contents
William Shakespeare	1. Act 1
	2. Act 1, Scene 1
	3. Act 1, Scene 2
	4. Act 1, Scene 3
	5. Act 1, Scene 4
	6. Act 1, Scene 5
	7. Act 2
	8. Act 2, Scene 1
	9. Act 2, Scene 2
	10. Act 3
	11. Act 3, Scene 1
	12. Act 3, Scene 2
	13. Act 3, Scene 3
	14. Act 3, Scene 4
	15. Act 4

**Figure 10:** WordSeer’s reading interface. If the document is subdivided into sections, these appear on the right as a table of contents.

For comparison and visualization of word prevalence, we have created the tool’s *newspaper-strip* visualization [1] (Figure 9(a)) Each play is represented as a long column. Within each column, small, colored horizontal blocks (corresponding to 10 sentences each) highlight the presence of a match. Hovering over a column (e.g. “Much Ado About Nothing” above) darkens it and displays the title. Hovering over a highlighted block displays the matching sentence.

Comparing the prevalence of different search words and grammatical relationships in different collections can give scholars thought-provoking results. For example, our scholar might compare the prevalence of references to *her* body parts to the prevalence of references to *her* relatives in the different collections. The visualization shows results for the *tragedies* collection are similar to the results for *comedies* (Figure 9(a)) but in *histories* (Figure 9(b)), an interesting pattern emerges. The visualization reveals

WordSeer supports quick, large-scale analysis through search and visualization, but in all cases maintains links back to the source text. Clicking on a result on the search page or on a highlighted block in the visualization opens up the reading interface (Figure 10). The full text of the document is loaded, and the system automatically scrolls to the relevant sentence, and highlights it.

For example, hovering over a blue or orange highlighted block in Figures 9(a) or 9(b) brings up a popup displaying the matching sentence. Clicking opens the reading interface to that point. The example scholar could thus quickly find interesting mentions of womens’ body parts or relatives and easily navigate to the sections of the plays containing those references.

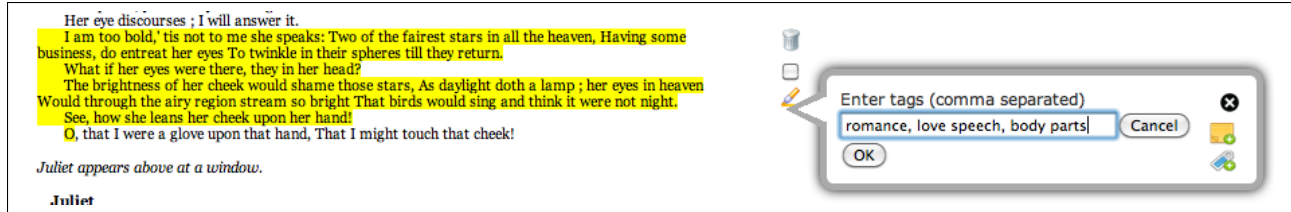


Figure 11: Highlighting text creates a “snippet”, to which tags and notes can be attached.’

## A.6 Annotation

WordSeer allows sections of text to be highlighted as “snippets” (Figure 11) to which notes and tags can be attached. This functionality is designed to help scholars keep track of interesting sections of text, and to organize them into conceptual categories.

In our example, a scholar might have formed the hypothesis that female body parts are predominantly referred to in romantic contexts. WordSeer’s annotation facilities can be used to tag speeches referring to female body parts by the topics the speeches seem to contain. In this case, it soon becomes apparent that many of the mentions are speeches by a lover.

## A.7 Exploration

For exploration of style and language, WordSeer uses computational linguistics to calculate *related words*. These words are either commonly used in similar contexts [4], or commonly used within a 10-sentence window of each other. Clicking on any word while reading brings up a small window showing related words.

In our example, the the related words for body-parts (Figure 12) help strengthen the hypothesis that female body part mentions are predominantly associated with romance. The popup shows that other body parts are frequently mentioned, along with love, fair, and sweet.

## A.8 Assembling Evidence

The process of constructing a final argument involves verifying hypotheses and collecting supporting evidence. In our example investigation, the hypothesis that female body part mentions are associated with romance can be tested by creating a final pair of categories. These are *not-love-stories*, in which love is not a major plot point, and *love-stories* for plays in which it is. When the plays are reorganized along these lines, the results are immediate.

In the *love-stories* collection (Figure 13(a)), we see *both* body parts *and* male relatives. By contrast, the *not-love-stories* collection (Figure 13(b)) shows *predominantly* male relatives, and hovering over the occurrences of body parts reveals a gloomy picture of *her* tear-stained cheeks and *her* sorrowful eyes.

Grammatical search results agree with the newspaper-strip visualizations and related words. We see more physical attributes *possessed-by her* in the in the *love* collection than in the *not-love* collection (Figure 14(a)–14(b)).

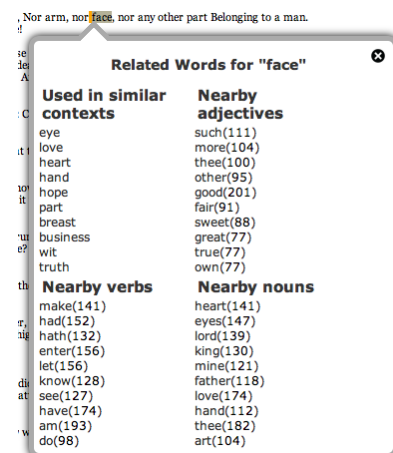
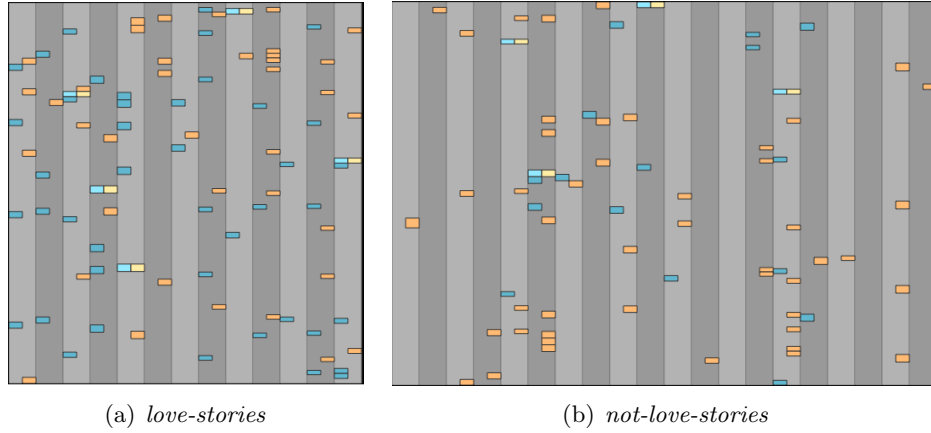
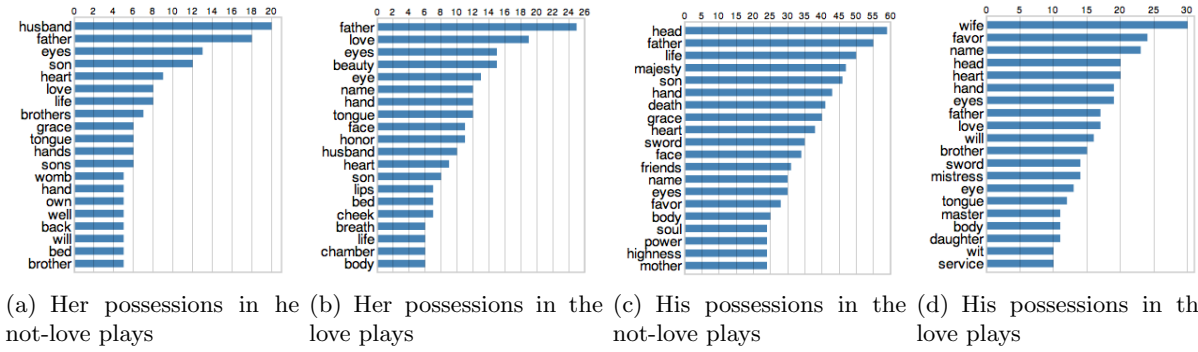


Figure 12: Related words for face



**Figure 13:** Comparing the *love-stories* and *not-love-stories* for the prevalence of body parts possessed-by her (blue) and relatives possessed-by her (orange).]



**Figure 14:** Comparison of grammatical search results for possessed-by his and possessed-by her

The grammatical search results show that the language around men changes as well (Figure 14(c)–14(d)). In the *not-love case*, the only woman to appear is **mother**, at number 20, but in the love case, **wife** takes first place, followed by **favor**. Compared to the physical language for women, these words have a more sentimental quality.

Thus, a scholar can gather evidence that suggests that, while a male-dominated picture of both men and women is always present, the physical aspects gain prominence for women in plays about love. For men, the more sentimental aspects come to the fore.



## B Current User Interface

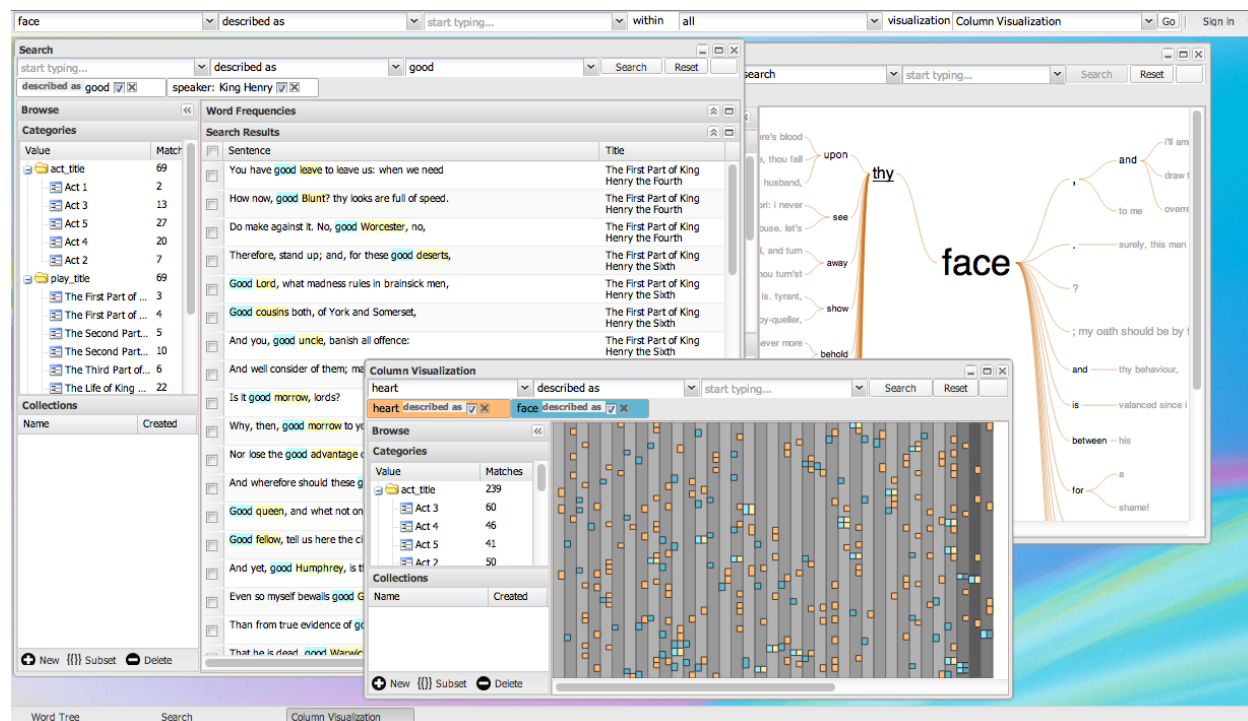
Previous versions of WordSeer put different visualizations and tools on different web pages. If a user was looking at a visualization, and then wanted to look at a different visualization of the same data, they would have to leave the first page, and go to another one. Then, they would have to type in the same search into the new visualization. To go back to the old one, they would have to use the browser’s back button.

This implementation had the advantage that a user could have multiple web pages open at once, but several disadvantages. Mainly, there were two problems: difficulty iterating, and difficulty with side-by-side comparisons.

By difficulty iterating, we refer to the fact that this implementation makes it difficult to do a set of quick, exploratory visualizations or searches. It is natural to want to refer back to older results, but they get removed from view as soon as the user goes to the next page.

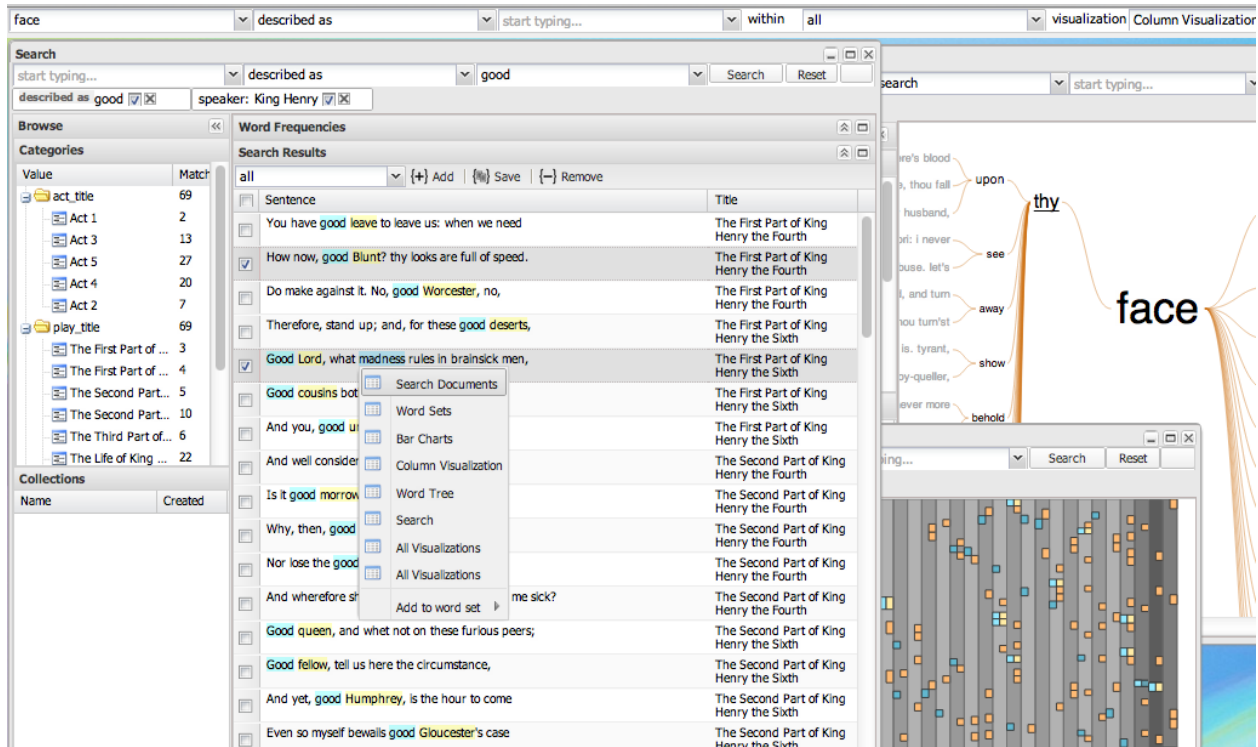
Second, by difficulty with side-by-side comparisons, we mean that the screen space taken up by an entire browser window is often so large that switching between two open visualizations in different windows was unwieldy. There was also the problem that searches entered in one window couldn’t easily be transferred into a different visualization in a different window.

We wanted an interface that would preserve the benefits of windows – multiple open at a time, with a flexible arrangement – but not at the cost of speedy iterations and easy comparisons. The ‘web desktop’ – so called because it resembles a desktop with windows and a task bar, but is actually a single, interactive web page within a web browser, seemed like the appropriate solution.



**Figure 15:** With the new interface, users can have multiple visualizations open at a time and be able to see both of them side by side. This alleviates the side-by-side comparison difficulty. Similarly, iterating is easier because all the windows are part of the same web page. They can share information.





**Figure 16:** In WordSeer’s new interface, new visualizations are easy to open based on the contents of existing visualizations. For example, one can click on any word in the search results and open up any number of visualizations of that word. In the old interface, such an interaction would have required three steps with multiple clicks in each step: opening up a new web browser window, navigating to WordSeer’s word tree, and typing in the search for the word.

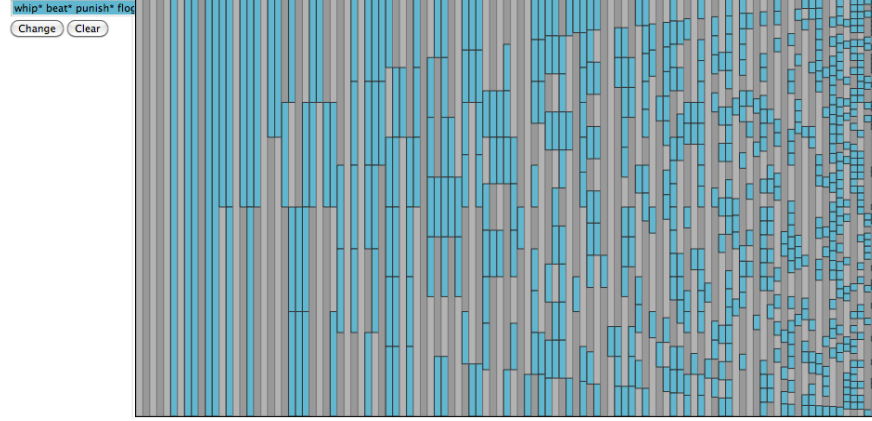
## C Case Study Report: Stereotypes in the North American Slave Narratives

In a 1984 paper entitled “I was born: Slave Narratives and their status as autobiography”, James Olney set out a number of narrative stereotypes that, he asserted, were “so early and firmly established that one can imagine a sort of master outline drawn from the great narratives and guiding the lesser ones”. His master plan was a numbered list that included:

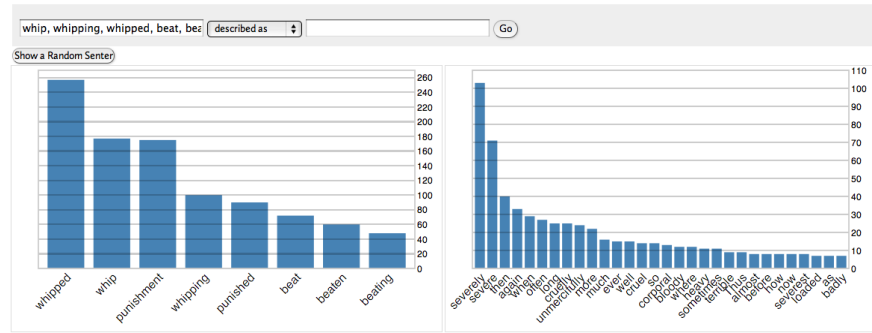
1. a first sentence beginning “I was born...”, then specifying a place, but not a date of birth;
2. a sketchy account of parentage often involving a white father;
3. a description of a cruel master, mistress, or overseer, details of first observed whipping and numerous subsequent whippings, with women very frequently the victims;

And continues on to:

8. account of slave auction, of families being destroyed, of distraught mothers clinging to their children as they are torn from them, of slave coffles being driven South;
9. description of patrols, of failed attempt(s) to escape, of pursuit my men and dogs;
10. description of successful attempt(s) to escape, lying by during the day, traveling by



(a) Widespread usage of words related to punishment.



(b) Widespread description of punishments in cruel terms.

**Figure 17:** Evidence of the “cruel punishments” stereotype.

night guided by the North Star, reception in a free state by Quakers who offer a lavish breakfast and much genial thee/thou conversation;

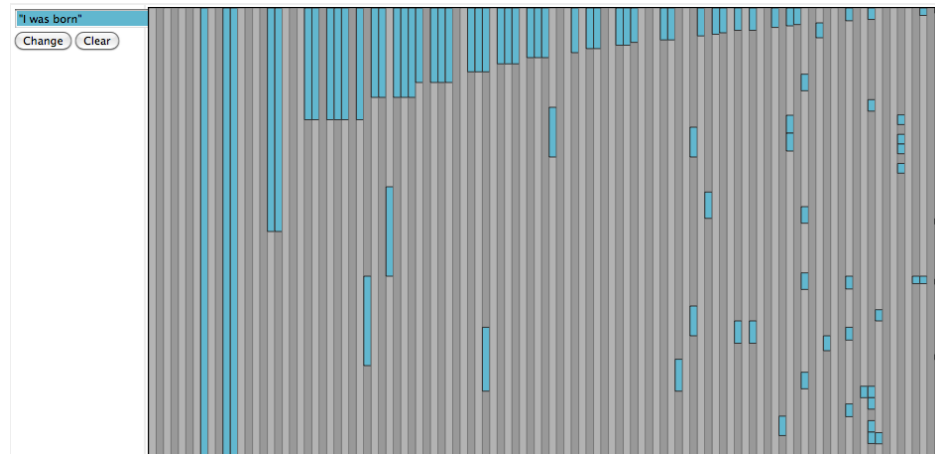
## C.1 Stereotypes

The first of the stereotypes we identified as common was the “cruel treatment” stereotype (3). Of the many listed by Olney, it was also the easiest to search for.

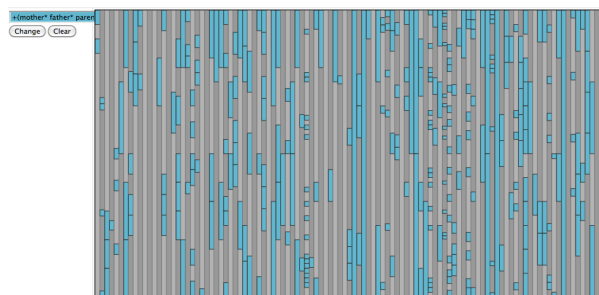
As shown in the newspaper-strip visualization in Figure 17(a), almost all the narratives had multiple occurrences of the simple keywords: “punish”, “beat”, or “whip”. Such events were not restricted to the few narratives Olney mentioned by name. Grammatical search (Figure 17(b)) reveals the other side of the stereotype. The numbers on the graph confirm the high prevalence revealed by the newspaper-column visualization, but the adjectives accompanying these actions paint a more complete picture: one of severe, cruel, and inhuman treatment.

Other similar stereotypes were stereotype 8, “families being destroyed” (Figure 19(a)), stereotypes 9–10 “escape” (Figure 19(b)), and stereotype 1, “I was born” (Figure 18). In all cases, searches for terms representing these stereotypes reveal widespread prevalence throughout the narratives.

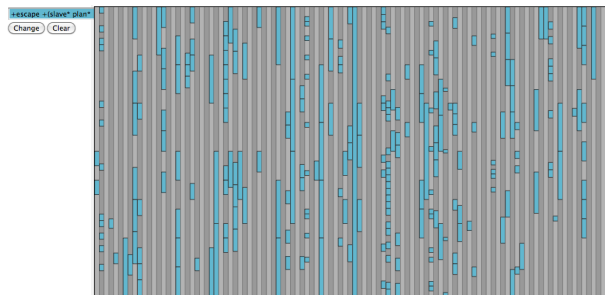
The stereotypes, however, did not seem to be ordered in any particular way. Except for “I was born” (Figure 18), which occurs predominantly at the beginnings of narratives, there does not seem to be any discernible ordering of stereotypes according to Olney’s numbering system.



**Figure 18:** The distribution of the stereotype “I was born”.

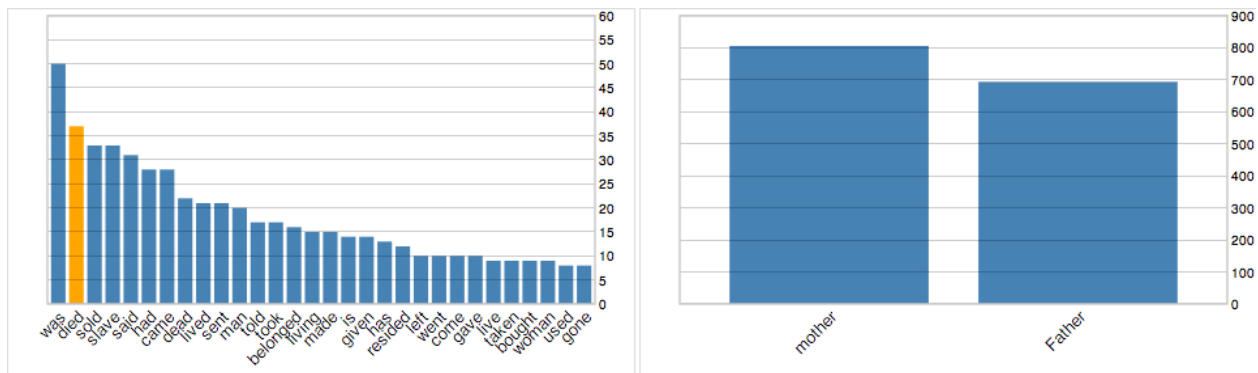


(a) Separation from parents.



(b) Escape from slavery.

**Figure 19:** Prevalent stereotypes.



**Figure 20:** The verb “died” is more common than “sold” in relation to “mother” and “father”.

## C.2 Missing stereotypes

While investigating the “families being destroyed” stereotype (8), we encountered what seemed to be a missing stereotype: the death of parents. In the list of verbs of which “mother” was the agent, “died” and “dead” were prominent. The same was true for “father”. In fact, when the two were combined, the prevalence of “died” was even greater than the prevalence of “sold” 20. If parents being sold was common enough to feature as a stereotype, this result seemed to indicate that their death should also qualify.

## C.3 Less prevalent stereotypes

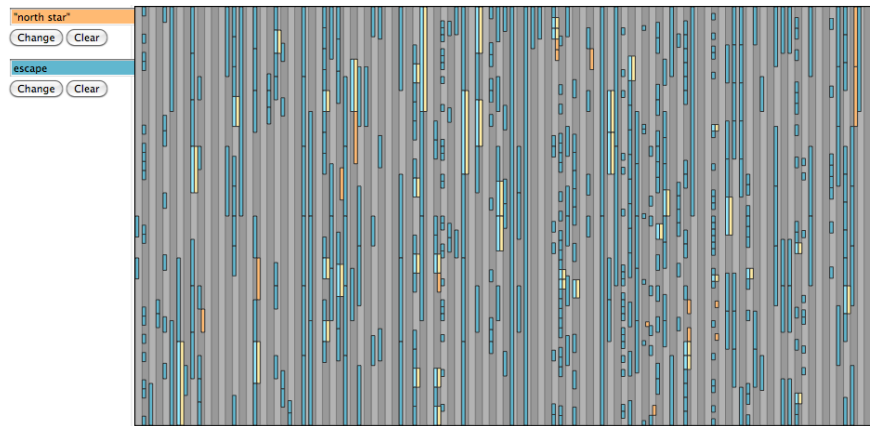
We were also able to identify at least two “stereotypes” that did not appear to be as prevalent in the collection as implied by Olney’s language: those of escapes guided by the north star, and of being received by Quakers.



**Figure 21:** Quakers: a less prevalent theme.

As shown in comparison with “I was born” (Figure 22), mentions of the north star do occur in some narratives. Upon investigation, we also found that almost all occurrences of the “north star” are indeed related to escapes. The remaining referred to a periodical called “North Star”. The accounts of escape, however, are far from being representative of the collection as a whole. Instead they are clustered around a few narratives that mention the north star multiple times. In particular, the title “Uncle Tom’s Companions: Or, Facts Stranger Than Fiction. A Supplement to Uncle Tom’s Cabin: Being Startling Incidents in the Lives of Celebrated Fugitive Slaves” contains 20 of the 86 references to the north star in the collection. A much more conservative “stereotype” around the north star seems to be indicated – whenever the north star is mentioned, it is always in relation to escape, but the converse is not true. In these narratives, escape is not necessarily guided by the north star.

Similarly, Quakers are mentioned in some, but not all narratives (Figure 21). When mentioned, they are always examples of kindness, and sympathy towards the abolitionist cause. Nevertheless, this convention is far less prevalent than the cruelty, escape, or separation stereotypes. Like the



**Figure 22:** Stereotypes compared. The visualization shows the occurrences of the word “escape” (blue) and the phrase “North Star” (yellow). The first is relatively more prevalent, and the two do not always occur in the same places.

north star, a more conservative restatement seems more appropriate. It is stereotypical to portray Quakers as sympathetic to slaves’ escapes, but not all escapes involve reception by Quakers.